

**Issued by the System Design Team on January 15, 2004**

Several URL addresses updated April 14, 2004

## **Assignments to Design Teams to Solve Metadata Problem**

**Problem:** Only 38% of currently visible layers in *The National Map* Catalog have metadata registered. Even our own holdings have only a 34% compliance rate. This is unacceptable. It's against our own policy, and has been a long-standing and growing embarrassment.

**Background:** The Standards and Directory Design teams have collaborated to collect the background material. The Standards Team reaffirmed that we have a clear policy already for metadata. This is an issue of compliance. The Directory Team summarized the current situation in a technical discussion paper "Geospatial Metadata in The National Map" (posted on internal partnerships web page). The entire issue was briefed to the IFORUM on Jan 12, 2004. Later that same day, the System Design Team discussed the issue and arrived at the following assignments to solve this problem:

### **Solution:**

1. While we act to correct this problem, we need to prevent this situation from getting any worse. Therefore, effective **immediately**, we are asking the Catalog Support Team (a part of the Directory Design Team) or others trained and certified to make and maintain Catalog entries to cease marking any content as publicly available unless it has URL-accessible metadata that meets the minimal FGDC and National Map metadata content standards.
2. The Standards Team is responsible for defining the minimum (normative) mandatory content for layer metadata.
3. We don't want to over-react by setting all non-compliant catalog content to "non-public" – this would confuse the public. However, we need to establish a reasonable timeframe to allow us to "catch-up" with this metadata population. Based on recommendations from the Standards Design Team who are familiar with what needs to be done, we are setting the due date for layer-level metadata population to May 14, 2004. Any layer not metadata compliant after this date will be marked "non-public" within the catalog.
4. Responsibility for populating: It is important to spread this workload to the groups who should have this as part of their normal job. This is not the responsibility of the Standards or Catalog Teams – those groups can both provide advice and training, but the operational responsibility for layer-level metadata population (seeing that it gets done) belongs to different groups:
  - a. For any of the eight layers held and managed by USGS, the responsibility for populating layer metadata resides with the specific Base-Data (theme) Design Team.

- b. For any of the eight layers provided by external partners, the responsibility for populating layer metadata resides with the MPO or C&R staff that developed the partnership.
  - c. For USGS content other than the eight layers, the responsibility for populating layer metadata resides with the data steward for the particular data holdings.
  - d. For non-USGS content, metadata will be accessible by means of the GOS /USGS directory synchronization process assigned to the Directory Design Team to develop.
- 5. The Catalog Operations Team will provide reports to each group summarizing the layers that have missing layer-level metadata.
- 6. To the extent this work is seen as new or unanticipated, the impact of this “catch-up” work should be reflected in the revised narratives being currently being prepared by each Center. If any group sees any major obstacles that would impact our goal of being 100% compliant with our own metadata policy by May 14, the System Design Team needs to be made aware of that compelling case as soon as possible.
- 7. The principle here is making this part of the job. If training is required, there are many sources of training available. FGDC has several training programs, and our own standards people can also assist in any needed training.
- 8. We also need to ensure that the technical connection between metadata generation engines employed by some of our internal holdings (NED and other SDDS holdings, for example) work well with the metadata functions in the viewer. We are asking the Leads of the Viewer Implementation and Catalog Teams to have technical discussions with Jean Paulson (for SDDS) and Steve Skelton (for GDA) to work out the automated interface specifications for generated layer metadata. This will be the mechanism we use to propagate source-file level metadata.

## FAQ

### **If I can't easily get the information from a partner, can I simply populate metadata fields with “unknown”?**

It depends on the particular field. Some fields permit an “unknown” value, while others do not. It is NOT an acceptable practice to simply value a field with the “unknown” value for the purpose of saving time or effort. There are certain mandatory fields that must be populated, and only certain fields can assume the value “unknown”. If you are unsure as to the minimum population rules, contact the Standards Design Team.

**Where can I get information on what things I need to worry about when registering content for *The National Map*?**

Specific technical requirements are described in the document “Registering Web Map Services in The National Map”, posted at [http://mciweb.er.usgs.gov/catalog/docs/wms\\_register.pdf](http://mciweb.er.usgs.gov/catalog/docs/wms_register.pdf). Send questions and requests for assistance to “GS-N-MCMC Catalog Support Team”.

**While the Standards Team is responsible for overarching standards (characteristics common to all themes; i.e., projections) and for some level of consistency, we have been deferring to the data theme leaders for content and accuracy standards.**

The Standards Team is responsible for the metadata field population rules. Population rules specify which fields must be populated, and the domain of acceptable values. The Eight Base Data Design Teams are responsible for defining minimum content specifications (inclusion criteria), which would include tighter ranges of acceptable values for certain metadata fields.

**Where exactly can I find the minimum metadata population rules?**

FGDC metadata standard: <http://www.fgdc.gov/metadata/constan.html>

The **workbook** for this standard: [http://www.fgdc.gov/metadata/meta\\_workbook.html](http://www.fgdc.gov/metadata/meta_workbook.html)

We recommend using the workbook rather than the standard for help. It is easier to use and contains a graphic representation of each section of the standard.

USGS metadata standard: <http://rockyweb.cr.usgs.gov/nmpstds/metastds.html>

Although the USGS document is referred to as a standard, the Standards for the Preparation of Digital Geospatial Metadata is actually the USGS implementation of the FGDC standard. This 'standard' includes the elements from the FGDC standard that we decided to apply to our products, and has sample metadata for each data type and each scale that they were produced (DLG, DEM, DOQ, and DRG). These documents may be used as a guide for the proper elements to include, however, the content of the elements (i.e. the responses to the element) will need to be altered for each dataset.

It should be noted that the Standards Design Team is currently in the process of recasting our current product-oriented metadata standards to a form more applicable to the web-services model. However, the content in the current product-oriented standard will be sufficient until the services-oriented version can be completed.

### **How did you arrive at the May 14 deadline?**

This is the recommendation from the Standards Team taking into account the following factors:

1. We want to correct this problem as soon as **practically** possible
2. The work to correct the problem is being spread over several Base Data teams and C&R/MPO offices
3. Some training will be required – this time is built in
4. Some time may be required to bring existing metadata into a URL-accessible form - this time is also built in
5. We are asking any group that can make a compelling case that the 14 May due date is unreasonable to make that case to the SDT as soon as possible. We don't want any surprises on May 15
6. We expect that communicating this decision/directive out to all affected teams should transpire NLT 1/14/04

### **Isn't this a policy issue?**

No – we have an established policy (see Standards team ppt briefing). This is a compliance issue.

### **I thought the System Design Team was supposed to work in design issues, not operational issues .....what gives????**

The Systems Design Team (and the individual Design Teams) have a dual responsibility – for design assignments, and for establishing operational procedures. The actual people that accomplish the work are in Centers and/or remote offices, but they are following the operational procedures defined by the design teams.

### **So, the actual assignment to fix this must ultimately come thru the Center Chiefs?**

Yes – that's why they will be cc'd in on this action request.

### **If this is part of the C&R job, it will take longer to add content – I thought we were to add as much content as we can as fast as we can.**

Yes, this is part of the job, and if that means completing the **whole job** takes longer, so be it. There may have been a time when we skimmed on metadata in favor of bulk content, but those days are gone (if they ever really consciously existed at all). The USGS has been a major supporter of metadata, and it simply sends the wrong message if we lack the commitment to follow-thru on this important aspect of geospatial data.

**Is the problem that we simply don't have any metadata, or that it exists, but the catalog doesn't know about it?**

We have a little bit of both, but from early reports, this appears to be mostly about getting the metadata we **do have** into a form the catalog can use. Generally, that means ensuring that the metadata (whether it's static, or dynamically generated) resides at some URL.

**What level of meta-data are we talking about here? Service, series, layer, product, file, or feature?**

The immediate problem concerns layer-level metadata. Let's solve that problem first. In *The National Map* implementation, each WMS layer should have exactly one metadata URL, which points to a file of metadata or to a metadata service. If a static file, the file must be in a browser-displayable format (HTML, plaintext, XML...).

We don't have a problem with service metadata - our current catalog population process ensures that service metadata is populated and correct. Access to source file and/or product metadata is covered in action #8. As for feature-level metadata, we'll attack that issue later. With the possible exception of NHD, very few of our holdings even have feature-level metadata. We will need to invest more in this area once we start moving into feature-oriented web services.